2005:12

LULEÅ UNIVERSITY OF TECHNOLOGY

RESEARCH REPORT

Improving Railway Punctuality by Maintenence

BIRRE NYSTRÖM
PETER SÖDERHOLM

Järnvägstekniskt centrum
Luleå tekniska universitet
Regnbågsallén, Porsön, Luleå
Postadress: S-971 87 Luleå
Telefon: 0920 - 49 10 00
Telefax: 0920 - 49 19 35
Hemsida: http://www.jvtc.luth.se

2005:12 • ISSN:1402-1528
ISRN: LTU - FR -- 05/12 -- SE

BANVERKET

Abstract

A railway has many different requirements, such as safety and punctuality. The purpose of this study was to explore and describe how maintenance can contribute to improved punctuality. In an industry branch that is becoming more open, as is the case in Norway where this study was undertaken, it is important to have clear requirements and performance measures.

This study has focused on railway stakeholders and their requirements regarding punctuality, the measures employed, and the traceability between requirements, goals, and measurements. Linking the end customer's punctuality requirements to reliability of the track is illustrated.

The requirements flow is presented as well as a discussion about its pros and cons. The results of the study can be applied in order to compare different designs of maintenance to fulfil punctuality requirements. The information flow and the measures currently used are not entirely suitable concerning feedback and usefulness for managing availability and thus maintenance.

Keywords: requirements engineering, stakeholders, punctuality, maintenance, railway

Sammanfattning

Många olika krav ställs på järnvägen, bland annat beträffande säkerhet och punktlighet. Syftet med denna studie är att undersöka och beskriva hur underhåll kan bidra till förbättrad punktlighet. I en bransch som blir öppnare och öppnare, vilket är fallet i Norge, där denna studie utfördes, är det viktigt att ha tydliga krav och prestationsmått.

Studien har fokuserat på järnvägens intressenter och deras krav beträffande punktlighet, vilka mått som används och sambandet mellan krav, mål och mått. Sambandet mellan kundens krav på punktlighet och banans funktionssäkerhet illustreras.

Flödet av krav presenteras och för- och nackdelar med dagens utformning diskuteras. Resultatet från undersökningen kan användas för att jämföra olika utformning av underhåll med avseende på hur väl krav på punktlighet uppfylls. Flödet av information och de mått som för närvarande används är inte helt lämpliga beträffande återkoppling och användbarhet för att styra driftsäkerhet och därmed underhåll.

Contents

1 Introduction

The railway industry is in a state of transition, with new stakeholders emerging and old ones finding themselves in a new environment. This puts a heavier emphasis on the need for systematic ways of handling requirements between stakeholders. This case study might contribute to this and be relevant to other logistical systems as well.

Punctuality – to be at the agreed place at the agreed time – is influenced by several stakeholders. The passengers or goods may not turn up when the train is bound to depart. The train operating company's (TOC) locomotives might break down, or the driver does not show up. Infrastructure malfunction, e.g. turnout failure, is another cause of non-punctuality. A train being late to a train meet at a single track might force a meeting train to wait at a siding, which in turn might cause delays to even more trains. The prioritising between different trains in such conflicts is handled by train traffic controllers, trying to minimise the spread of such disruptions. So, we see that customer's behaviour, TOC's equipment and personnel, infrastructure maintenance as well as planning and scheduling of train traffic influence punctuality.

The restructuring of Norwegian railway began 1989 with the partition of Norwegian State Railway (NSB) in divisions. Since 1996 Jernbaneverket (JBV) is a separate body responsible for infrastructure management and NSB is a TOC (http://www.jbv.no (04-04-01)).

Norwegian/Swedish terms are put in brackets.

2 Aim and goal

There are a lot of different requirements on the railway. Examples of important requirements are safety and punctuality. The goal of this study is to explore, describe and analyse how requirements on punctuality and related terms are handled in railway organisations. The study aims to explain how maintenance can contribute to improved punctuality.

The main questions are

- How are goals on punctuality formulated and deployed in organisations using/used by the railway?
- How do punctuality goals affect the maintenance of investigated railways?

3 Objects of investigation

The phenomena to be studied are railway stakeholders'

- perception of the word punctuality and related terms
- requirements in (which requirements from whom)
- requirements out (which requirements to whom)
- feedback/control (how the stakeholders receive feedback respectively give feedback).

Both the present state and the desired state, according to the stakeholders, are investigated.

3.1 The stakeholders in this study and their relations

The stakeholders in this study stand in quite similar relations to each other as corresponding organisations of other European countries, where the railway has been reorganised during the 90's. The organisations investigated were chosen in order to facilitate the description and analysis of a chain of requirements and allow for comparisons between countries. Therefore, the Norwegian Post was chosen as the freight customer. See Figure 1, which also shows the other main stakeholders in the value chain for delivering a letter.

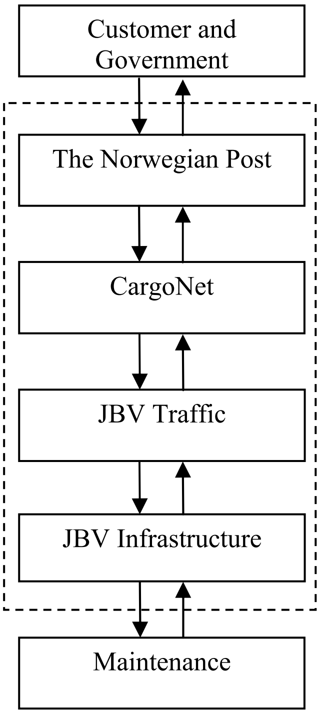

Figure 1. The value chain in transporting a letter by rail. In this model, requirements are symbolised by downward arrows and feedback, while measuring the requirement by upward arrows. Of course there are other factors than stakeholders' requirements that also play a part in the control of the processes, but they are not studied here, as important processes' external requirements are the main objects of this study. The scope of the study is given within the dashed rectangle.

4 Method

Here we describe how the study is carried out. We begin discussing the approach to stakeholders and then discuss the view of the investigators on the role of the stakeholders and interviewees.

4.1 Stakeholders

To identify stakeholders with requirements on punctuality a methodology influenced by process mapping was selected, as described by Juran (1992) and Sharp et al (1999). In order to collect stakeholder requirements on punctuality semi-structured interviews was selected as an appropriate methodology. The collected data was deployed with the aid of Matrix Diagrams, in a manner influenced by Quality Function Deployment (QFD), see Akao (1992). Based on the requirements on punctuality, quantitative measurements were identified in order to establish a model that can be applied for the comparison of different solutions in the design of maintenance, in order to improve the punctuality.

In this case study, interviews and document studies were used in order to find out how job tasks are carried out. To begin with, a few top managers in railway stakeholder organisations and blue collar workers were interviewed about their conception of, and role in, punctuality and maintenance. With the help of results from these interviews, key professionals were identified and interviewed, in order to get the full picture. The focus was to get an understanding of which requirements they received, in which form and who from, and how they handled these requirements, e.g. passing them on. Regulations, handbooks and other documents were studied when interviewees referred to such documents.

4.2 The stakeholders and interviewees

The stakeholders are organisations, or parts of organisations, or individuals, but the interviewees are persons. We investigate the roles of the persons as stakeholders, e.g. the role as quality manager.

Two very different views of investigations involving humans as sources for information are used by the sociologists Weber respectively Durkheim, as summarised in Hughes et al (2003):

Weber	Durkheim
Subjective	Objective
Social action	Social facts
Motive for action	External forces

According to Weber's view, one has to understand the view of the individual studied, not just facts, as they might be seen different by different persons. The individual's motive for action is important to Weber, while Durkheim has a positivist view and argues that the factual circumstances of the individual in the society e.g. limitations, are forces which cause the individual to behave in a specific way. For example, the time and space perceptions of the individual are modelled after the social organisation, according to Durkheim. Weber argued that individuals act for rational reasons. These rational reasons might be different from individual to individual, e.g. for a captain of a ship it might be rational to be the last one leaving his ship. In this case we have so called value rationality.

We acknowledge that both views have their pros and cons, but that practical considerations in this study advocate certain views. As the number of interviewees in each stakeholder role is low, the detailed quantification of the impact on external forces is difficult. We have to trust the motives for action given by the interviewees, at least if this motive is corroborated by several persons in similar functions, and the persons giving requirements to, or receiving requirements from, the interviewee.

5 Basic concepts and definitions

There are different definitions of requirements. According to ISO/IEC STD 15288 (2002) "stakeholder requirements are expressed in terms of the needs, wants, desires, expectations and perceived constraints of identified stakeholders." In the above definition stakeholder requirements include, but are not limited to, the needs and requirements imposed by society, the constraints imposed by an acquiring organisation and the capabilities and limiting characteristics of operator staff. Davis (1993) defines a requirement as "a user need or necessary feature, function, or attribute of a system that can be sensed from a position external to that system". Characteristics of and descriptions of good requirements are in Appendix 2. Dimensions of requirements.

A stakeholder can be defined as (ISO/IEC STD 15288, 2002): "an interested part having a right, share or claim in the system or in its possession of characteristics that meets that party's needs and/or expectations". In this definition stakeholders include, but are not limited to, users, supporters, developers, producers, trainers, maintainers, disposers, acquirer and supplier organisations, regulatory bodies and members of the society.

According to Davis & Leffingwell (1996), Requirements Management is a systematic approach to elicit, organise, document, and manage both the initial and the changing requirements of a system. A principal result of this work is the development of one or more requirements specifications, which define and document the complete external behaviour of the system to be constructed. Davis & Leffingwell (1996) consider Requirements Management as partly a management process and partly an engineering discipline, and states that therefore it can be used effectively to manage both technical complexity and requirements on the system.

Kotonya & Sommerville (1998) describe Requirements Management as the managing of changes in the requirements, and the relationship between requirements on a system. Requirements Management verifies that it is technically and economically possible to perform proposed changes. If the change applies to a specific requirement it is important to check which other requirements that can be affected. This requires that links between requirements, the sources of requirements and the system design must be documented, i.e. traceability information.

Requirements Management includes methodologies and tools to establish and execute a formal procedure to collect, verify, consider, and study how changes affect the system (Kotonya & Sommerville, 1998). Thereby, Requirements Management can be seen as a way to manage mainly four activities by the feedback from these activities (Bohner & Arnold, 1996). The first step is to study the proposed change and make decisions about necessary changes and appropriate actions, based on wanted and unwanted effects of the change. After that, the changes must be specified and designed. The changes must then be executed based on the specification, e.g. through document changes. Finally, the performed changes must be studied to see if they meet the new requirements, and that the system meets existing requirements.

There are different definitions of punctuality, which coexist, and the difference in the views of different stakeholders play a crucial role, as will be detailed later in the stakeholder interviews. The authors will use the word punctuality to describe the times of arrival and departure at places important to stakeholders, while the term delay will be used to describe the deviation from schedule in-between these places.

6 Scope and outline of work

The major view of Requirements Management in this report will be as the management of changes in the requirements, and the relationship between requirements on a system. The maintenance of infrastructure is in focus. This means that the traffic is taken for granted, and e.g. trains must not be moved in order to get more time for track maintenance.

First, we identify stakeholders and identify the way they express their requirements and to whom. Then, we analyse their way of expressing their requirements and discuss alternative ways of expressing them. So, the present state will be described as well as the desired state.

6.1 Information looked for

We strive to capture the informants' uses and views of the word punctuality and related terms. The meaning of the word to the informant, as well as which qualities of punctuality they experience in their work, is studied. So, in this part, the study has as hermeneutical as well as phenomenological perspective. (Although the word respondent is used for interviewees expressing their subjective thoughts according to Olsson & Sörensen (2001), we use the word informant, as most of the information is factual.). Other requirements than punctuality are asked for analogously in order to get an understanding of performance measures and data used by informants in order to improve. The importance of this relies on the underlying assumption that, what is measured gets attention and people try to improve it. The “What to achieve” of one stakeholder transforms to the “How to achieve” of the next stakeholder (Kaplan & Norton, 1996) and we want to investigate requirements that go between stakeholders. We are aware that this approach excludes requirements between stakeholders within the organisation (which in this study turned out to make up a significant proportion of total number of requirements) as well as internal efficiency requirements, e.g. concerning amount of capital goods and ROI (return on investment).

6.2 Information gathering

In order to be capture possible classifications of punctuality and requirements, literature has been studied. A focus group is used to fill in missing classifications as a preparation for the individual interviews.

6.2.1 Focus group

A focus group is used in order to fill in missed classifications of punctuality and requirements. Focus groups also have the advantage of “forcing” the participants to motivate their opinions. In this study, a so-called PULS-group, cooperative group of personnel from railway manager’s organisation, train traffic control and TOCs was used. (The interview was conducted as detailed in Appendix 1. Focus group. In principle, the same steps are carried out when performing individual interviews.)

6.2.2 Interviews

Individual interviews are carried out. Face-to-face interviews were preferred, as it gives more facetted as well as a higher volume of information. They were followed up by telephone calls to complement missing issues.

The interviews are structured around punctuality perception and aspects of requirements. Roughly, they follow the group interview scheme except that the informant is asked to tell more about his/her professional background in order to give the interviewer a better background understanding and less is in writing (point 3 of Appendix 1. Focus group). The interviewer funnels down the issues, as illustrated in Figure 2.

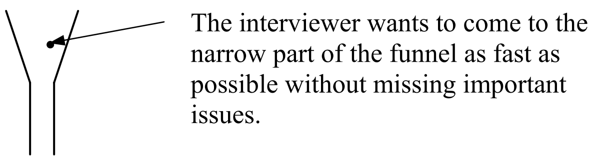

Figure 2. Interview funnel. The interviewer starts questioning broadly concerning the interviewee perception of punctuality and pins down the classifications used by the interviewee. The same process starts again when asking questions concerning requirements, where requirements affecting punctuality are discussed in detail. To avoid lengthy background discussions in the funnel’s top, the interviewer has to gain confidence by e.g. showing that he is knowledgeable in the field (Rubin & Rubin, 1995).

Identification of punctuality classifications (the document Dimensions of punctuality) used by the informant are strived for. This is done by asking for the informant’s view of the word punctuality. It is asked where he/she encounters punctuality in his/her stakeholder role. The informant is encouraged to give examples of punctuality and related terms. Straight-forward why-questions are avoided. The Dimensions of punctuality document works as a checklist when interviewing. The interviewer has this document so the informant sees it. It is used by the interviewer to ask questions (and possibly read definitions) to the informant, if the informant's answers do not clearly favour one classification for the other.

Identification of requirements is done by asking for requirements on the informant. The questioning concerns requirements in general and the interviewer uses the funnel method to pin down requirements concerning punctuality. Then the feedback the informant gives is asked for. The requirements the informant has on other stakeholders are then asked for along with the feedback he/she gets. A form based on Appendix 2. Dimensions of requirements and Appendix 3. Dimensions of punctuality, functions as checklist when interviewing.

6.2.3 Information verification

In order to verify information given by the informants, documents, especially when referred to by the informants, are studied. These are statistics on punctuality and maintenance, regulations etc. Another way is the comparison between one informant's requirements out and corresponding requirements from other informants.

6.3 Analysis

Analysis is performed listening to the interview tapes, classifying interviewees' classifications of punctuality and requirements.

Transcription is performed of all material, as recommended by ten Have (1999), in order to avoid bias. This is done by first listening through the recording in order to be able to more easily recognise the words the next time. Transcription is performed on a word by word basis, simply marking stress with underline and ignoring a systematic description of e.g. breathing. Peculiarities in the material are marked by comments.

ten Have (1999) recommends to follow the steps in the analysis as follows.

1. Turn-taking
    a. Identify type of action: question, answer, objection…
    b. Identify technical terms
    c. Investigate how turn-taking supports certain understandings of terms
2. Identify sequences i.e. how one thing leads to another.
3. Identify repairs. I.e. identify repetition due to not audible or not understood utterances.
4. Turn-construction/design is described by ten Have (1999) as the most unclear of the steps. He considers it to involve discussions on what's behind utterances.

To us, step 4 is the most important. The steps 1-2 helped us to perform it. Step 3 was considered unimportant in the interviews. The steps were carried out for the focus group interview as well as the individual interviews.

After transcription, data fragments are coded. Then fragments that have the same code are retrieved, in this way facilitating the comparison of similar data fragments as described by Rubin & Rubin (1995). Concepts, represented by words or sentences, ideas and stories are looked for. The coding strives to concentrate the meaning of the arguments. The underlying paradigm relates to grounded theory, according to ten Have (1999), who refers to Glaser & Strauss (1967).

The analysis is performed after each interview to gain a deeper understanding until next interview, especially in order to be able to investigate the similarity of requirements from one stakeholder to another.

7 Results

Here we report on the findings, stakeholder by stakeholder. We begin from the top of the value chain in Figure 1, with the Norwegian Post, and work downwards the supply chain. Main requirements and requirements considered important to this study are described here.

7.1 Results from the focus group

From the PULS-group, no additional dimensions of punctuality were found than the ones already known (see appendix Dimensions of punctuality). One interesting point was the group's emphasis of punctuality as relative to an agreement (or expectation), not relative to timetable. A multitude of punctuality related expressions were used[1] by the participants. The perceptions of punctuality of the TOCs were more oriented towards customers' punctuality, while other organisations related asset reliability and causes for unpunctuality. A positive measure, thus avoiding delay minutes or the like, was wanted by several participants. Furthermore, the coding of the focus group interview revealed the themes information in advance concerning unpunctuality, agreements (timetable) versus expectations. The fact that agreements are made in a chain was briefly discussed.

From the discussion about requirements we excerpt the point of one TOC with shipping companies as customers. As the shipping companies requires the TOC to be able to fulfil their weekly demand of transportation, the TOC makes weekly measurements to see how many requests it turns down due to lack of capacity. The measure "number of requests turned down this week" is said to quantify lack of capacity, the disadvantage being that customers that have learned about the lack of capacity will not come back the week in question.

The group discussion added a dimension to the classifications of requirements. To have proactive (as opposed to reactive) requirements was considered important. E.g., when not entirely so, was safety inspections made by the railway inspectorate. To learn who should act how, in order to improve, was also considered important. (However, we do not consider this to be features of the requirement itself, but of the organisation's knowledge.) Furthermore, the group gave disadvantages of slightly fewer measuring methods than it gave advantages. So the interviewer learned that he might have to ask explicitly for disadvantages when carrying out the individual interviews.

7.2 The Norwegian Post

The mission of the Norwegian Post is to deliver integrated, value-adding communication and logistics solutions through physical and electronic networks. Nearly all Norwegians have their mail delivered six days a week (http://www.posten.no (04-02-26)). Here we study physical letters and packages.

7.2.1 Organisation

The Norwegian Post is organised in five divisions

- Communication. Includes internal development functions and e.g. stamps.

[1] In no specific order: punctuality, in time agreed, according to decided timetable, right time, expected time, precision of delivery, satisfied-customer-index (broader term?), punctual, delay minutes (opposite?), right time-minutes, delay, quality of traffic, arriving in time, and in time.

- Logistics. Includes e.g. market function and the partly owned Pan Nordic Logistics.
- Consumer. Includes post offices.
- Distribution net consists of Distribution (local distribution in the proximity of a terminal (regionpostgång)), Production (means sorting) and Transport (transport between terminals (hovedpostgang)).
- ErgoGroup AS. Includes e.g. administrative functions and electronic services.

(http://www.posten.no (04-02-26))

To us, the Distribution net is interesting, as it is the division directly involved in the transport of mail. There are about 20 terminals in Norway. (Senior consultant, Norwegian Post)

7.2.2 Transport chain and punctuality and quality measurement of the Norwegian Post

Figure 3 and Figure 4 show the logistic flow of packages. This model of the logistic chain aims to show material flow between sender and addressee of a package, as well as recording point for passage of package or vehicle, in order to analyse times and explain punctuality. (Efficiency, e.g. utilisation rate of vehicles, is not described by this model.)

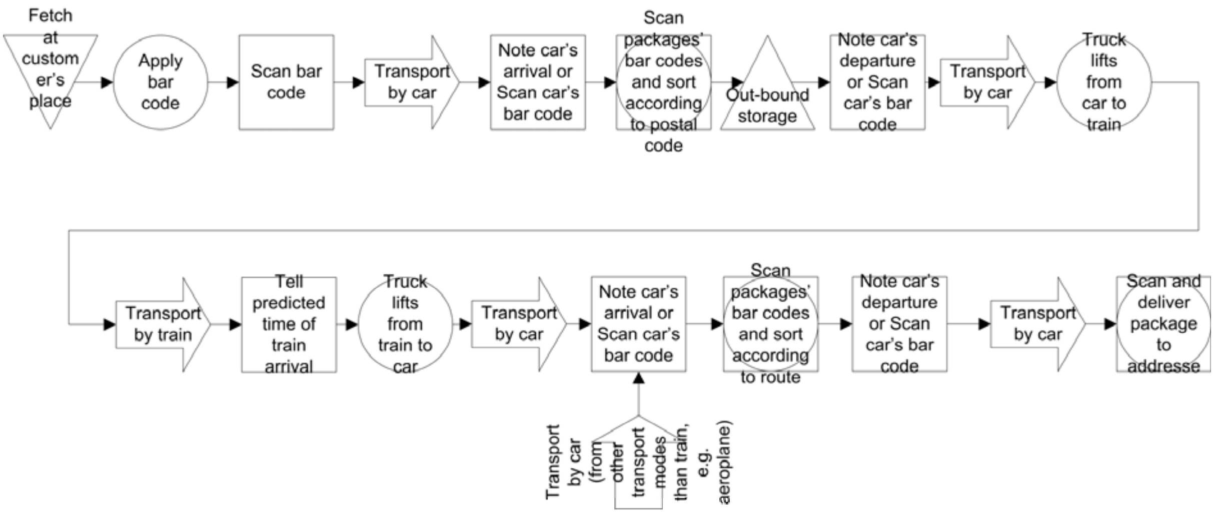

Figure 3. The logistic chain for packages, when using train as means of inter-terminal transportation. Small containers (växelbehållare) are used for post transportation. The arrows represent transport, circles operations and squares control points (kontrollpunkter), where the position of the goods or the vehicle is recorded. Car transport from terminal to railway station is triggered by messages. The car has to be at the station well before train departure, a time which varies (1.5 h in Trondheim according to a truck driver) on order to get the train loaded. The arrow Transport by train includes shunting of wagons to and from train. The last step, transport to the addressee, might be carried out by car, bicycle or foot.

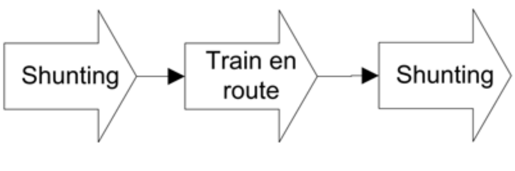

Figure 4. Detailing operation Transport by train.

7.2.2.1 Track and trace technology

Recording arrival and departure time of vehicles transporting post to and from terminals by pencil and paper is still in use in some places. It has shown to be not entirely reliable. The entrepreneur, eager to avoid having to pay delay penalties to the Post, might object to allegations of being late. He might support his claim using car's tachometer recordings, these probably being more reliable than man-made recordings. As a bar code is applied on each package when handled in, it allows "track and trace" of the individual package by scanning the bar code when loading, unloading and sorting packages. Drivers of vehicles are to register their departure and arrival to post terminals by registering who they are by choosing their identity from a menu of bar codes. The next step in technology is radio frequency identification (RFID). This allows for non-contact reading by having e.g. vehicles, letters, packages equipped with RF tags pass through a door frame equipped with an antenna that emits radio signals to activate the tag and read data from it. The elimination of the possibility of drivers accidentally registering as another driver is an argument for this while the investment cost is an argument against an implementation now, according to one informant (Senior consultant, Norwegian Post).

There is no way of tracking individual letters directly, as they have no bar code applied. But the vehicle is tracked, as for packages.

There is no data collected on when the car actually unloads to the railway respectively loads from the railway, instead car's movement is controlled by train's expected time of arrival and time of car's departure from and arrival to terminal are recorded. So prediction of railway arrival and departure time is very important. As the factual arrival and departure times of the train are recorded, one cannot by this method tell whether the car was late or the loading/unloading procedure caused a certain delay. It is thus difficult to tell which to improve in case this is an important delay cause.

In the transport chain, delays (førsinkelser) might strike both departure and arrival. Whether transports are delayed or not is followed up. Although the number of minutes delayed is measured, it is not possible to follow up, according to Transport manager Trondheim, Norwegian Post. Regularity (regularitet) is a term used for the fraction (letters, train, aeroplane) being on time compared to time plan (departure and arrival). Left-behinds (genligg) are post left behind at a place, exceeding plan.

7.2.3 Requirements to the Norwegian Post according to the Post

The Norwegian ministry of communication and transport (Samferdseldepartementet) requires the Post to deliver 85% of first class letters (A-brev) within Norway to the addressee the day after posted in the post box. This quality of service target is measured as national average per calendar month (Senior consultant, Norwegian Post). An independent firm of auditors measures this percentage by sending test letters (packages are not sent). Addressees notified beforehand, record when letters arrives to them. Some letters have RFID, allowing them to be traced through the transport chain. This approach allows for analysis of separate relations, e.g. from Oslo to Bodø, although a very detailed analysis is not possible due to the limited number of letters.

Fast packages (expresspakke) are also required to be delivered overnight, although this is not a governmental requirement. Other products have longer delivery times, e.g. four days for a package product aimed at private persons (Norgespakke). The product manager (produkteier) within the Post identifies the market requirements for different products and acts as an internal customer.

There are penalties for late delivery of some products, giving the Post customer the fee back when delayed.

7.2.4 Requirements to the Norwegian Post according to Customers

The requirements above were found by asking the Post. In this study, we take these requirements as is, i.e. we do not ask the Post customers what they want, as it is considered to be out of the scope of this study. Furthermore, the requirements might vary considerably from customer to customer. The Post uses internal product owners which learn customers' requirements and, in turn, buy matching products from the producers in the Post.

7.2.5 Requirements from the Norwegian Post

All transportation modes are utilised by the Post; aeroplane, train, truck, car and boat. Transportation is carried out by its own personnel as well as by entrepreneurs.
Long distances make the use of aeroplane a necessity to deliver overnight. The Post has no dedicated post aeroplanes and uses ordinary scheduled passenger flights (SAS). Train post containers are transported by CargoNet. Information on train delay is needed in order to make the Post able to reach an informed decision whether to distribute the post already arrived in time and hand out the post from train later, or to delay the distribution of all post.

The Post strives for penalties which reflect the marginal cost of the Post for the delay in question (Senior consultant, Norwegian Post). As an example, the penalty for the flight company for not transporting a kg of post is higher for the last flight of the day than for earlier flights. There are no flight delay penalties for the airline. For road transports, the penalty steps are 30 min, 1 h and 2 h. For train, there is one step; 30 minutes.

7.2.6 Discussion on requirements to and from and within the Norwegian Post

About half of the requirements given by the Post lack method of measurement. Of such requirements to the Post, several concern timeliness[2], not punctuality, of fetching or delivery from/to the Post customers and the Post terminals. Of such requirements from the Post, several concern timeliness respectively information. The Post requires from its business customers to allow fetching of post in a certain time span (or several time spans, if there are much to collect) in order for the Post's cars to travel the shortest route at the shortest time with as full cars as possible; this is today measured as cost/km and income/km, in the future it will be complemented by kg/car. We think that the cost/km is about the same for fetching late packages (16.00) as for the more expensive 9.00-packages, and that the higher costs for the 9.00-packages are in intra-terminal transports, (income-cost)/km being a somewhat crude measure e.g. when comparing similar vehicles travelling different routes; then kg/car is better, as it shows capacity utilisation and margin more directly, facilitating evaluation and planning of route and vehicle choice. A requirement internal to the Post is that the Transport division should be on time to the terminals, but not deliver all post at the same time in order for the terminals to get a uniform workload. Delay information should be passed from CargoNet to

[2] By timeliness (läglighet) we understand to what extent a journey/transport happens when convenient to a stakeholder.

the Post. The timeliness is hard to quantify and the delay information is said to have improved considerably recently, to a satisfactory level.

The consequences of delay are measured as left-behinds, i.e. post not being forwarded in the transport chain according to plan (genligg), measured as (e.g.) numbers of packets by route from and postal code to. Regarding letters, this is a good measure, as letters not being forwarded one day are forwarded at the latest the day after, most probably delivered. Packages being delivered directly to the addressee might, on the other hand, be delivered later the same day. I.e. packages that are promised to be delivered before 9.00 might be delivered at the same time as 16.00-packages. As packages are individually bar coded, we therefore conclude that there is a good data basis to measure punctuality of delivery for letters as well as packages.

The amount of disappeared post is measured by the number of customer complaints. The number of delayed packets might be measured by postage reimbursed, as some products have such a delivery time guarantee. As some customers will not complain, we consider these statistics probably have to be complemented by sample tests. The test letters are useful especially in measuring punctuality of local post transport (lokalpostgången), as intra-terminal transports (hovedpostgangens) are traced but local post transports' activities of emptying letterboxes and delivering letters are not recorded.

The non-following up of length of delay might make it harder to adapt planned time to the actual transportation times. As there might be several transports close in time, delay time might be known quite well anyway.

We remark that the manual measurements of arrival and departure of vehicles makes analysis more difficult, due to the introduced errors.

7.3 CargoNet

CargoNet offers transport of containers as well as entire system trains (freight trains configured for certain, large industrial customers). Containers might be moved to and from other transport modes; truck and boat. CargoNet is owned by Norwegian NSB and Swedish Green Cargo (http://www.cargonet.no (04-06-10)).

7.3.1 Organisation

CargoNet have 13 terminals in Norway. The number of employees is 980, including e.g. locomotive engineers, terminal personnel for loading and unloading and administrative personnel (http://www.cargonet.no (04-06-10)). CargoNet perform maintenance of their rolling stock themselves.

7.3.2 Punctuality and quality measurement in CargoNet

Punctuality is measured as the percentage of trains which have begun unloading (alternatively delivery time at the customer's) within one minute after time agreed with the customer. The punctuality is measured for each relation and delays are given in time intervals (Quality manager, CargoNet). In the case of a train being split into two shorter ones taking different routes, concerning arrival punctuality this counts as two trains, see Figure 5. CargoNet's own goal is to reach more than 90% punctuality. The unloading time varies and is about half an hour after the scheduled arrival of the train.

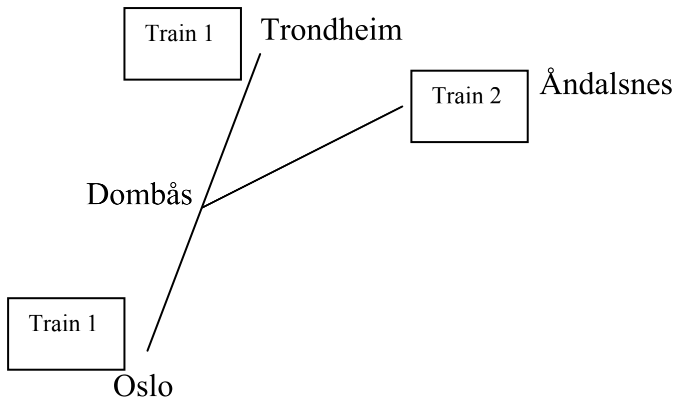

Figure 5. A train from Oslo is divided into two trains and then one of the trains is assigned a new train-number.

In the transport chain, CargoNet measures punctuality at departure, underway and at destination, underway punctuality meaning to what extent the transport time is according to schedule (regardless of departure time) (Manager operations centre, CargoNet).

7.3.2.1 Track and trace technology

The locomotive driver notes train times on paper; when collecting train from workshop as well as time for train being connected to the wagons and departure and arrival times. He handles over the documents to personnel at the terminal, who feed it into CargoNet’s computer system. CargoNet consider this to be resource consuming but it gives the ability to check to what extent each function in the transport chain sticks to the agreed times. We comment that according to our experience from other trains, data is probably unreliable and hence requires filtering out erroneous data when doing the analysis.

7.3.3 Requirements to CargoNet according to CargoNet

Customers demand faster transport, later handle-in and earlier collection, as well as punctual delivery. This requirement is met on a customer by customer level (the average train speed or similar was not used by e.g. the interviewed Manager operations centre, CargoNet). Meetings with customers are scheduled once or twice a year; issues include punctuality and damages to goods. The number of complaints is recorded. According to CargoNet, this open way gives suggestions to improvement, but the selection is limited. The flexibility of having extra train paths available (behovsruter) has been reduced in recent years, as there are more stakeholders on the track nowadays. Train paths are free of charge; fee is paid for actual traffic.

There are penalties for late trains for some customers, e.g. the Post, but not in general.

Information on delayed trains is nowadays sent to a dedicated e-mail address for each customer. Easy-to-understand bills are requested by the customers, and is said to have improved recently from being a frequent cause of complaint (Quality manager, CargoNet).

7.3.4 Requirements from CargoNet

The requirements to JBV concern agreed train departure and arrival times as well as having traffic restrictions due to planned maintenance being announced well in advance, more than a few weeks.

Another important requirement is for JBV to give information concerning on how long time a repair of an occurred failure is prognosticated to require, as well as to have short repair times. CargoNet has no overall statistics concerning these matters, but there are JBV – CargoNet meetings where occurred failures are discussed.

7.3.5 Discussion on requirements to and from CargoNet

Some of the requirements given by CargoNet lack obvious method of measurement. One such requirement to CargoNet concerns timeliness of train routes. This is also true concerning CargoNet's requirement on track time. Furthermore, we were not able to find a clear definition of force majeur i.e. when the requirements are dropped.

We consider information to be of utmost importance to the TOC, in order for him to be able to plan measures; inform his customers and decide on alternative means of transportation. The performance regime has only one step for trains, 30 minutes, compared to three for road transport. We think that the Post's marginal cost will not be covered for considerably longer delays, and that this regime will not give incentives to reduce causes to delays.

7.4 Jernbaneverket (JBV)

Jernbaneverket's (JBV; Norwegian National Rail Administration) organisation is seen in Figure 6. JBV is to manage the overarching requirements to the railway sector from the Norwegian ministry of communication and transport (Samferdseldepartementet) namely, co-ordination between traffic modes, investments and maintenance, efficiency, regional policy, environment and safety. There are no quantified targets directly concerning these goals (Maintenance director, JBV). There are governmental decisions concerning how to reach the goals, e.g. install ATC (Automatic Train Control) on specific sections to increase safety (http://odin/dep.no (04-05-21)).

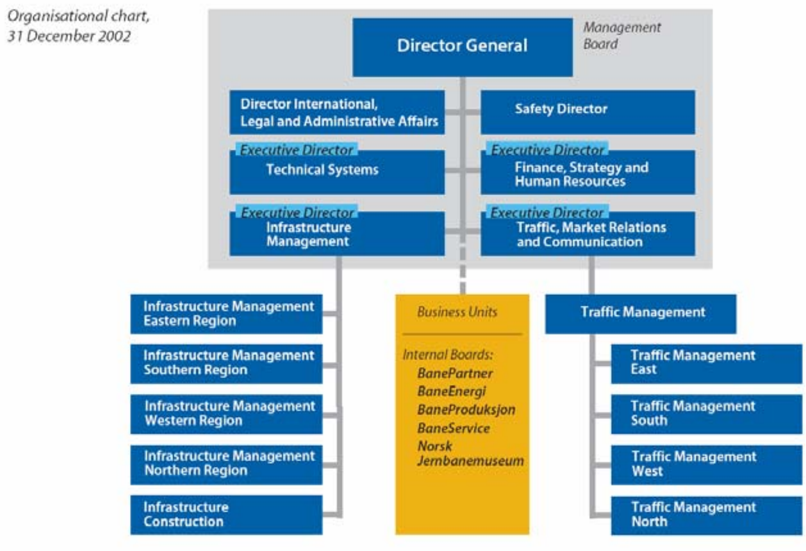

Figure 6. JBV's organisation 2002.

7.5 Jernbaneverket Traffic Division

7.5.1 Organisation

Jernbaneverket's traffic division is divided into three regions (recently reduced from four), spanning the same areas as the corresponding regions of the infrastructure division.

Each region has a traffic manager (trafikksjef) and each region has two to four track managers (banesjef), managing one or some lines. The track managers are subordinates of the regional manager (regiondirektør) (http://www.jbv.no (04-06-16)). Maintenance personnel are subordinated to one track manager, but perform work also in other track managers' parts of the region (Track manager, JBV).

7.5.2 Punctuality and quality measurement

Punctuality is measured as percentage of trains arriving to final destination on schedule. On schedule means, for local trains and medium-distance trains, within 3 minutes after scheduled time, 5 minutes for long-distance or freight trains. Train departures/arrivals before schedule are considered as punctual in the statistics. Time loss (tidstap) due to planned track works is the other indicator found in JBV's reporting (JBV Punktlighetsrapport 2003). This is a theoretical measure on how much reduced speed for a track section (due to e.g. dåligt spårläge) prolongs the transport time of trains passing the section (it does not consider knock-on-effects, i.e. trains being delayed by delayed trains). Regularity (regularitet) is a term used for the fraction of trains not cancelled. What a cancelled train exactly is, was somewhat unclear to us from some of the informants. A cancelled train according to JBV is a train that does not travel its entire route (Analyst, Traffic control, JBV).

7.5.2.1 Track and trace technology

Where there is centralised traffic control train passage events are recorded automatically, but there are a lot of stations between main stations that relies on train dispatchers to manually register train arrivals and departures. The number of train delay minutes is aggregated nationally and divided among responsible bodies, but there is not detailed statistics published.

7.5.3 Requirements to JBV Traffic according to JBV Traffic and comparison

Basic requirements come from the Norwegian Railway Inspectorate (SJT, Statens Jernbanetillsyn), Norwegian Labour Inspection and the employees concerning railway safety, safety at work and work environment in general. Revisions by the Railway Inspectorate and registered number of incidents give a measure of to what extent regulations are followed and incidents occur (Traffic director,JBV).

Train operators require 90% of trains on time to end station (95% for the Gardemoen airport train) and JBV Traffic has this as an internal requirement for different categories of trains; short distance, regional traffic and for each freight train product (container express, system train,…) (Jernbaneverket Virksomhetsplan 2004-2007. Trafikkområde Øst).

Other requirements include passengers requiring information on train traffic according to the service statement of JBV (serviserklæringen) on railway station information, namely appropriate timetable, delay information at the latest when the train was scheduled and about platform change. JBV measures the number of passenger complaints for each line and has noticed that the number of complaints rises when information equipment is out of order. The disadvantage is that number of complaints measures only the top of the iceberg. We think that e.g. the number of complaints might be lowered if equipment is not repaired for a long time, as passengers become fed up. Another, more objective complementary measure, might be the number of failures.

TOCs also require good information to their operations centre, locomotive drivers and customers. To what extent this is fulfilled is told from the complaints.

Operators require fair prioritising between different trains, namely that trains on time go first when there are track access conflicts and that prioritisation is performed roughly from airport train to freight train. (Traffic director, JBV). Operators require timeliness, enabling later delivery of freight to departing freight trains (Traffic director, JBV)

TOCs require robust and effective schedule, i.e. the schedule to be able to cope with naturally occurring deviations. This is enabled by allowances and margins. This is measured by the punctuality of the traffic, a measure known by all stakeholders although it is important to complement it with track utilisation percentage (Train traffic planner, JBV). We remark that this is a post measure; proposed schedules are not evaluated for their robustness.

JBV infrastructure division requires train traffic planners at JBV to utilise the capacity of new assets by shortened transport time or increased frequency. One indicator is slack time or cost/benefit for the asset. This is an easy-to-understand indicator, although the TOCs wish for a changed station stopping pattern might corrupt a straight-forward slack time indicator (Train traffic planner, JBV).

JBV infrastructure division requires access to track to carry out maintenance. The measure here is e.g. cost/km track. The disadvantage of this measure is that it does not consider the impact of maintenance for TOCs (Train traffic planner, JBV). We remark that such a measure must be complemented with the state of the equipment before and after the studied period or be used over a long time to even out the effects of reinvestments.

The requirements above were found by asking JBV Traffic. We now compare these requirements to the ones given by the ones having requirements to JBV Traffic.

7.5.4 Requirements from JBV Traffic

JBV traffic division wants to learn when TOCs (train operating company) cancel trains, as it is the duty of the JBV to inform passengers on this.

JBV traffic division wants the locomotive driver to communicate operations problems to the traffic control first, then to his company. The TOC is also required to fulfil requirements concerning type of rolling stock (speed etc) which is seen in the TIOS computer system (Traffic director, JBV).

The JBV traffic division requires delays to planned reduced speed on track section (tidstap) to be at the most 1 min/100 km. This might be performed any way, e.g. working another time of day, use another method, work faster and coordinate track works. The requirement is easy to administrate, as every track manager (banesjef) knows how large a time he might use.

The JBV traffic division has no direct requirements to the infrastructure manager concerning time to repair or alert level, it is left to the track manager. How to measure the uptime of the infrastructure is currently a subject of discussion. Track availability is required, e.g. in order to allow high train speed immediately after track work ballast cleaning is to be followed by tamping, thus being a process requirement (Traffic director, JBV). We conclude that currently we have not found any quantification of uptime requirements, making comparisons between e.g. regions difficult as well as e.g. estimating the consequences of a future traffic increase.

JBV traffic division requires from TOCs to tell which type of rolling stock they intend to use and prospective train schedules (rutebestillinger), possible to carry out. Achievement is measured as train punctuality (Train traffic planner, JBV).

JBV traffic division requires track managers (banesjefer) to have worked out train plan data and track works planning. This is measured by deviation from agreed time lost because of temporary speed restrictions (tapstid) and planned track work time (vita tider) exceeds into train path. This is easy measurable although statistics are lacking. (There are no economic transactions scheduling – track managing involved here.) (Train traffic planner, JBV).

7.5.5 Discussion on requirements to and from JBV Traffic

A lot of requirements are said to have punctual trains as the measure.

7.6 Jernbaneverket Infrastructure Division

7.6.1 Organisation

Jernbaneverket's three infrastructure regions correspond one-to-one to the traffic regions. Each infrastructure region has two to four track managers (banesjef), each managing one or a few lines. The track managers are subordinates of the regional manager (regiondirektør) (http://www.jbv.no (04-06-16)). Maintenance personnel are subordinated to one track manager, but perform work also in other track managers' parts of the region (Track manager, JBV). There is also a division for investments (utbygging) (http://www.jbv.no (04-06-16)).

7.6.2 Requirements to JBV infrastructure according to JBV infrastructure

The basic requirement from TOCs is time on track supplemented with axle load etc.

There is no general requirement to JBV infrastructure concerning time until repair begun or repair time, and these times are not measured (according to Maintenance director, JBV). With the Flytoget airport train, JBV has an agreement that repair personnel are to be in place within half an hour, although there is no penalty if not. Train delays due to prolonged track works are measured. JBV's track managers have to pay substitute buses for TOCs when trains are stopped by not agreed track works. There is no time limit. Money might possibly go in the other direction if TOC hinders track works. (Track manager, JBV).

7.6.3 Requirements to JBV infrastructure according to JBV infrastructure and comparison

Basic requirements come from the Norwegian Railway Inspectorate (SJT, Statens Jernbanetillsyn), Norwegian Labour Inspection and the employees concerning railway safety, safety at work and work environment in general. Revisions by the Railway Inspectorate and registered number of incidents give a measure of to what extent regulations are followed and incidents happens (Traffic director, JBV). (And, of course, to what extent they are reported.)

7.6.4 Requirements from JBV infrastructure

Requirements from JBV infrastructure HQ goes to the regions of JBV infrastructure.

The region is to calculate product cost indicators (produktnyckeltal), e.g. cost of one metre tamping and indicators for track quality and production for each track manager section (banenummer) (Maintenance director, JBV). There are about ten goals in total (see Jernbaneverket Vedlikeholdshåndbok Rev 0.1 1B-Ve 29.01.04). The goals for the track itself include track quality number, requiring it to be more than 95 for all but the lowest prioritised tracks. The goals also include contact wire failures, allowing less than 0.2 failures per 10 km of track causing train delays and goal for the number of signalling failures. Goals for other infrastructure failures, e.g. turnout failure, are not given. The target for delays due to planned track works and unanticipated speed restrictions is 1 min per 100 km track. Finally, the delays[3] due to infrastructure failure are to be less than 50 hours per million train-kms. The arrival delay is to average less than 2.5 minutes[4].We see that the goals are formulated as quality (track quality number etc, connected to travel comfort and risk of derailment), numbers of failures per km track and the consequences of failures. If we consider the delays due to infrastructure as the top goal, one understands that other infrastructure equipment has to cause less delay than “what is left”. How much is left? It depends on downtime due to failure, traffic volume and mix, not only the number of failures. From a top-down perspective, this way of breaking the goals into smaller ones, is not good, as it is an easy job to have a low number of failures, e.g. contact wire failures, if you have a small traffic volume. On the other hand, comparing one’s number of failures with other, similar, tracks, provides a motivating as well as realistic target to strive for. The targets are set with safety first, most obviously including the requirement zero lateral buckling of rails due to temperature variations.

Maintenance actions are reported to the Maximo maintenance system after being carried out.

Every year, the region is to formulate a renewal plan spanning 10 years and compare current state to the earlier planned state for each track section (banenummer). In this renewal plan, maintenance strategy is to be described and cost/benefit analysis of projects calculated. (Maintenance director, JBV and see e.g. Jernbaneverket notat 03/5556 SHK 023). We have not heard of any post-implementation evaluation of measures carried out in JBV.

The regions are to keep track of preventive maintenance backlog (eftersläp). There is currently no systematic way of describing the backlog (Maintenance director, JBV).

There are no targets for delays or other disturbances caused by rolling stock.

7.6.5 Discussion on requirements to and from JBV infrastructure

Failures to infrastructure are classified into either track, contact wire or signalling. The latter ones both have goals concerning their numbers, but track in total has not. If one considers delay time to be of importance, it is strange, as delay time for failures of either track or signalling has about the same mean and standard deviation, while contact wire has longer delay time and quite many very long times[5], which we consider renders the number of contact wire failures being a fuzzy goal.

[3] Delays mean a train is lagging schedule. A train delayed a certain number of minutes at a station underway receives delay minutes, even if it arrives at its final destination on schedule (Analyst, Traffic control, JBV).
[4] Jernbaneverket Virksomhetsplan 2004-2007. Infrastruktur RØ. “Hoved-, Gjøvik- og Gardemobanen” 08.03.2004. We remark that this goal is fulfilled, at least in the report’s eastern region.

[5] For freight in Norway 2003 the count, mean and standard deviation of infrastructure-caused train delay minutes were as follows. Track (86, 36.4, 60.0), Contact wire (223, 118.0, 153.6), Signalling (751, 26.2, 36.4) (extracted from JBV database).

Of track failures, track quality number, track misalignments (vindskevheter), gauge widening (sporutvidelser), broken rail, lateral buckling of rail due to temperature variations and number of places with crushed ballast (plaskepartier) have goals. So we see that these goals concern one or several of ride comfort, the state of the track, as well as direct accident and delay-causing circumstances. A lead indicator is the number of crushed ballast points.

Inside JBV, we remark that the causes in the database are not entirely correctly classified, e.g. maintenance-induced delays (e.g. due to track machine breakdown). This is important to learn, so one can handle risks (including for delays) when planning maintenance.

8 Results summarised

Here we summarise the results. The most obvious chain of requirements regarding punctuality concerns delivery time, see Table 1. The greyed boxes hold requirements as given by the informants. In white boxes are the same requirements, formulated “the other way around”. In the middle column we have placed the requirement formulated in a “positive” way, in the rightmost the same requirement formulated in a “negative” way. As seen from Table 1, only one of the four stakeholders (JBV infra division) formulates its requirement in a negative way.

We see that the definitions of punctual differ slightly. Concerning CargoNet – JBV Traffic, CargoNet has smaller tolerance, but this time includes the time to shunt the wagons to unloading. CargoNet considers trains unloading for customers in their punctuality statistics, while JBV considers trains to end station. The last step, from punctuality to reliability of contact wire, gives rise to some comments and objections.

Table 1. A chain of requirements. Inside the bold outlined boxes are requirements as found in the case study. The other requirements are formulated by the authors.

Stakeholder	Requirement to stakeholder	Equivalent requirement
Norwegian Post	Minimum 85% of Priority Class letters delivered over night (national average per calendar month).	Maximum 15% of Priority Class letters handed in on time still belong to ‘remaining post’ in the morning after (national average per calendar month).
CargoNet	90% of trains on time (< 1 min) to station and unloading has begun.	Train departure, underway transport or unloading delay totals ≥ 1 min for at most 10% of trains.
JBV – traffic div.	90% of trains on time (< 5 min) to end station.	10% of trains ≥ 5 min late to end station.
JBV – infra div.	99.999 % of functioning trains pass 10 km of track without contact wire problems per year (approximate).	Contact wire < 0.2 failures giving operations difficulties per year and 10 km of track.

9 Discussion and conclusions

There are several ways to measure the performance of a process

1. measure to the process
2. measure within the process
3. measure performance (output)
4. Measure a consequence of performance, e.g. customer satisfaction (i.e. letting customers assess the output of the production process)

Underway punctuality might be seen as internal to the process, as the customer's concern is departure and arrival times (process output). The stakeholders used punctuality to final destination as the punctuality, although CargoNet also sees departure punctuality as important in order to be able to reach the final destination in time. Input is often money used.

The Norwegian Post uses internal product managers to learn customer requirements, CargoNet uses enquires, customer meetings and the number of complaints. JBV traffic division measure complaints concerning passenger information as well as complaints from freight carriers. Media discussions of passenger traffic performance are also a performance measure. JBV infra division does not has the same variety of customers and measures process parameters impact to each customer.

The Norwegian Post measures deviations from schedule left-behinds for each node and route, CargoNet transporting the post measures punctual trains, JBV traffic also does so, using a slightly different definition of punctual. JBV infrastructure measures primarily number of failures. Strangely, up time (opptid) of track is not defined within JBV; it is needed to provide the link between failure and delays. To measure repair time will be necessary in order to calculate up time.

The word regularity was interpreted in different ways by the informants, meaning to which extent post is punctual or not, or trains are de facto run, or to which extent trains are ready for departure at scheduled time. We conclude that ambiguous use of the word might be confusing in customer-supplier relations, although each interpretation is line with the broad definition of regularity in the Norwegian Z-016 standard on regularity and reliability on the shelf of Norway.

The important term cancelled train means the same, a train not travelling its planned route, to both NSB and JBV, although some informants were uncertain of the exact meaning of the phrase. The number of cancelled trains is measured and goals are given for the number of cancellations caused by JBV. Cancelling a train, substituting it with (slow-going) buses along the whole or a part of its route, will show up in the statistics as a cancelled train. There is no delay registered for the buses, so cancelling trains reduces delays. We conclude that the praxis, found in some reports, to also declare number of cancelled trains, should be encouraged.

Few of the indicators were quantified, even in cases where they lend themselves to quantification. Quantifying requirements is especially important when dealing with requirements that are in opposition to one another. An example is the wish to have a robust timetable as well as small allowances and margins. Another example is planned delays due to planned maintenance actions which might be in opposition to the wish to have track utilisation. How high goal to set, is still a question whether one subscribes to the goal view (e.g. cut by half from today's state) or optimisation view (achieve the most economical state).

The track and trace information was considered to be good by the TOC CargoNet, but CargoNet considers the prognoses of the JBV on failure remediation time not to be entirely satisfying. Although we were unable to find hard data, here is probably improvement potential for maintainers to give good estimations of repair time.

The temporal dimension of indicators show that most indicators are lag indicators, i.e. they tell the past, making analysis of prospective actions more difficult. Punctual arrival to final destination is an example, number of injuries another. In order to be proactive, one needs lead indicators. An example of a lead indicator, i.e. an indicator that predicts the future is ballast crushed points, which predicts future accidents and reduced track speed, causing delays. A set of more forward-looking indicators should be aimed at.

10 References

10.1 Publications

Akao, Y. (1992): Quality function deployment: integrating customer requirements into product design. Productivity Press, Cambridge, Mass.

ANSI/IEEE STD 830

Bohner, S.A. & Arnold, R.S. (1996): Software change impact analysis. IEEE Computer Society Press, Los Alamitos, California.

Davis, A.M. (1993): Software requirements: objects, functions, and states. Prentice-Hall International, Englewood Cliffs, New Jersey.

Davis, A.M. & Leffingwell, D. (1996): Requirements Management in Medical Device Development. Medical Device & Diagnostic Industry Magazine, 18, 100-116.

Hughes J. A. et al (2003): Understanding classical sociology – Marx Weber Durkheim. Second ed. SAGE, London. ISBN 0-7619-5467-8.

IEEE STD 830

ISO/IEC STD 15288 (2002): Systems Engineering - System Life Cycle Processes. ISO/IEC.

Jernbaneverket notat 03/5556 SHK 023 Vedlikeholdsstrategi 2005-2014.

Jernbaneverket Vedlikeholdshåndbok Rev 0.1 1B-Ve 29.01.04

Jernbaneverket Virksomhetsplan 2004-2007. Infrastruktur RØ. "Hoved-, Gjøvik- og Gardemobanen" 08.03.2004

Juran, J.M. (1992): Juran on Quality by Design. The Free Press, New York.

Kaplan, R. S. & Norton, D. P. (1996): The balanced scorecard: translating strategy into action. Harvard Business School Press, Boston, Massachusetts c1996. ISBN 0-87584-651-3.

Kotonya, G. & Sommerville, I. (1998): Requirements engineering - processes and techniques. John Wiley, Chichester.

Norwegian Technology Standards Institution (1998): NORSOK Standard. Regularity management & reliability technology. Z-016. Rev. 1, December 1998. Oslo, Norway.

Olsson, H. & Sörensen, S. (2001): Forskningsprocessen: kvalitativa och kvantitativa perspektiv. Liber, Stockholm. ISBN 91-47-04958-8.

Rubin, H. J. & Rubin, I. S. (1995): Qualitative interviewing: the art of hearing data. Sage Publications, Thousand Oaks, Calif. ISBN 0-8039-5095-0.

Sharp, H., Finkelstein, A., & Galal, G. (1999): Stakeholder Identification in the Requirements Engineering Process. In 10th International Workshop on Database and Expert Systems Applications, pp. 387-391. Institute of Electrical and Electronics Engineers, Florence, Italy.

ten Have, P. (1999): Doing conversation analysis: a practical guide. Sage, London 1999. ISBN 0-7619-5586-0.

Yin, R.K. (1994): Case study research: design and methods. Second Edition. Sage, Thousand Oaks, California.

10.2 Informants

Personal communication, spring 2004.

Analyst, Traffic control, Jernbaneverket.

Maintenance director, Jernbaneverket.

Traffic director, Jernbaneverket.

Train traffic planner, Jernbaneverket.

Track manager, Jernbaneverket.

Senior consultant, Norwegian Post.

Transport manager Trondheim, Norwegian Post.

Manager operations centre, CargoNet.

Quality manager, CargoNet.

Appendix 1. Focus group interview

The focus group interview is conducted in the following way:

1. The moderator (the researcher) introduces himself and the aim of the exercise. (Estimated time: 3 min.)
2. The group members introduce themselves (useful when later identifying persons on tape). (Estimated time: 2 min.)
3. A blank paper is handled out and the informants are asked to fill in their answer to the questions What is punctuality to you? Do you use any other word associated with punctuality? The participants are allowed to write during 5 minutes. It is emphasised that no opinion is right or wrong, that no decisions are taken on the grounds of their opinions and that the purpose is to generate an extensive, good description of the field. After writing, the moderator allows everyone to read their texts. The group members are encouraged to supplement each other's views. The moderator listens for opinions on each of the dimensions in Dimensions of punctuality and additional dimensions. The use of (hypothetic) stories is anticipated to play an important role. (Estimated time: max 20 minutes)
4. The moderator gives a brief presentation of the next task for the group. It concerns requirements in the railway. The term requirement is illustrated by a confectioner having requirements to fulfil as well as having requirements on others. (5 minutes)
5. The group members fill in forms concerning requirements: what requirements, from and to whom they have requirements. Also how feedback is given respectively received and their opinion on the form of feedback. (10 minutes?)
6. The group members are asked to fill in + or – for each requirement, indicating how it affects punctuality.
7. The group discusses the requirements written down one by one. The moderator strives to classify each requirement according to dimensions of requirements identified so far and listens for additional dimensions. The group members are encouraged to comment each other and give examples from their experiences. (15 minutes?).
8. The moderator thanks the group members for participating in the study.

Appendix 2. Dimensions of requirements

Types of requirement could be described along several dimensions. We here describe the ones seen in literature, by our own thoughts and explored during the group interview.

A requirement could be *exact*, giving the exact number of items required, the exact maximum cost etc. As it is often impossible to be exact, e.g. concerning point in time, a *tolerance* specifying minimum and maximum allowed value should be used. A requirement of the type “in 90% of the cases…” or alike, we call a *probabilistic* requirement.

A requirement could be either *functional* or *non-functional*. Functional requirements tell the function of the system, i.e. both input and output and their relation. Non-functional requirements tell the attributes of the system, including level of productivity, safety, maintainability, compliance to standard, etc. (Level of unpunctuality is thus a non-functional requirement.)

A requirement could be regarded as *mandatory*, *desirable* or *advantageous*. This prioritisation of requirements could be done using other classifications as well, with less or more than three classes, the extreme is to put all requirements in a total order.

A requirement could be *proactive* or *reactive*. A proactive requirement indicates that measures should be taken in order to attain a certain desirable consequence (or avoid an undesirable consequence). To be able to manage one has to have lead indicators, that is indicators telling what will happen in the future. One example is the amount of money invested in new production equipment or personnel education. Lag indicator is the opposite, telling what has happened. Most accounting figures, e.g. rate of return on capital for last fiscal year, are lag indicators.

For a certain requirement, it might be easy or difficult to tell who is responsible for fulfilling it. Furthermore, it might be easy or difficult to tell what to do in order to fulfil it. Anyway, these are not features of the requirement itself, but to the organisation and its capability to establish cause-effect relationships.

E.g. consider requirements for a street light bulb. The (basic) functional requirement is that is should emit light. The non-functional requirements specify how much light. The minimum amount of light and its quality might be stipulated by law, i.e. mandatory requirements (also non-functional), as well as by higher standards of the street owner, i.e. desirable requirements (also non-functional). A lifetime requirement (also non-functional) might be expressed in probabilistic terms, e.g. that is must be a 90% probability that any bulb last more than one year.

Characteristics of and descriptions of good requirements. According to ANSI/IEEE STD 830 (pp.11-13) and IEEE STD 830 (pp. 4-8).

Characteristics	Description
Correct	The requirement shall state something that the system shall meet.
Unambiguous	The requirement shall have only one interpretation. As a minimum, this requires that each characteristic of the final system be described using a single unique term.
Complete	The requirement shall be both determined and significant.
Consistent	The requirement shall not conflict with other requirements.
Ranked for Importance and/or Stability	The requirement shall be ranked for importance and/or stability if it has an identifier to indicate its importance or stability.
Verifiable	The requirement shall have some cost-effective methodology with which a human or machine can check that the system meets the requirement.
Modifiable	The requirement shall be possible to change easily, completely, and consistently without affecting the structure and style of the requirement specification.
Traceable	The requirement shall have a clear development from its origin to its furthest development.

Appendix 3. Dimensions of punctuality

Here, we briefly define punctuality and related terms. (The Swedish terms are put in brackets.) One of the goals with the interviews is to find out which definitions and principles the interviewees use.

Definition. By *timely* (läglig) we mean that the journey/transport occurs when convenient. ■

From this definition, departure as well as arrival times are included (and hence the travel time). The schedule, as well as deviations from schedule, affects the timeliness (läglighet). Such deviations might be considered bad by the customer (e.g. too late arrival) or good (e.g. an expected departure deviation from schedule gives shorter change time and consequently shorter total travel time).

Definition. By *timeliness* (läglighet) we mean to what extent the journey/transport occurs when convenient. ■

Remark: one might tell the stakeholder before timeliness, e.g. *customer timeliness*.

Definition. By *punctual* (punktlig) is meant that an event takes place when agreed. ■

Note: When persons agree to meet at a certain time, we consider them punctual if all of them show up in time. By event we also mean that an agreed state has been reached – e.g. that the painter has made the house painted to the agreed time, or that the train has reached a certain location at the agreed time. Defining punctual in this way leaves the value of punctual undefined when the agreement is unclear. Punctual is a logic variable; an event is either punctual or not punctual. The antonym is *unpunctual*.

Definition. By *punctuality* (punktlighet) is meant to what extent an event takes place when agreed. Punctuality of an event is described by tuplets of the form
<Stakeholder, Location, Event, TimeOfEventAgreed, TimeOfEventOccurred>,
where

Stakeholder → Passenger | GoodsCustomer | TrainOperatingCompany
| InfrastructureMaintainer | RollingStockMaintainer | …
Location → AnyLocation
Event → Arrival | Departure | Passage **of** Passenger | Goods | Vehicle
TimeOfEventAgreed → DateAndTime
TimeOfEventOccurred → DateAndTime | Cancelled
■

Note: Only events (i.e. occurrences at an instant), not continuous episodes (happening over periods in time), might have the punctuality property. The same event might be described by several tuplets, as there might be e.g. several stakeholders of one event. A continuous movement, e.g. a train moving along the track, might be partly described by a finite number of punctuality tuplets. The Passage Event means that the train passes by without stopping. Partially cancelled trains are described by Cancelled for each location, as it is the non-event of arrival or departure that is described, not the act of cancelling the train. A drawback of the above description is that it generally does not manage events that are wished not to happen.

Examples. Two tuplets describing events occurring at the same time:
<Passenger Karlsson, Trondheim Central, Arrival of Passenger Karlsson, 2004-03-12 13.00, 2004-03-12 13.02>
<TrainOperatingCompany NSB, Trondheim Central, Arrival of Train 70, 2004-03-12 13.00, 2004-03-12 13.02>

The tuplet set
<_, Trondheim Central, Arrival of Train 70, _, _>
contains every arrival of train 70 at Trondheim Central at least once. There will be one tuplet for each stakeholder of every train 70 arrival at Trondheim central.

An important decision when telling punctuality (and timeliness) is the choice of *system boundaries*. The choice is between the alternatives

- Passengers'/goods' punctuality
    - Each person's/piece of goods' punctuality is concerned. It might be approximated e.g. by assuming a certain number of passengers on a certain relation. The system boundary is nonetheless the same.
- Vehicles' punctuality
    - Each vehicle's punctuality is included. Trains or wagons might be considered. It might be measured for certain relations.
- A certain system's reliability. Here, (un)reliability numbers of a system with importance to (un)punctuality are considered. E.g. number of passenger display failures, wagons' failures or point failures.

We also have to choose between

- Cross-mode, i.e. concerning a chain of transports on a certain relation or
- Unimodal, i.e. concerning only one mode of transport e.g. train transport.

Another choice is between *discrete* and *continuous punctuality*. By discrete punctuality is meant that a person (or goods, train) is either considered punctual or not. On the contrary, continuous allows the punctuality to be described in a continuous manner. A discrete definition of punctuality defines a threshold (5 minutes is the most common in Europe). Trains arriving or departing later than this threshold from the schedule are considered unpunctual. A discrete punctuality scheme would tell e.g. that 90% of trains are punctual to their terminal station. Many schemes are a combination of discrete and continuous, e.g. the Japanese 1 minute threshold (trains less than 1 minute after schedule are considered punctual; discrete component) and summing up all unpunctuality minutes (continuous component). There might be several discrete steps. (Strictly, as most railways truncate or round-off punctuality data to whole minutes, the unpunctuality measure is discrete, not continuous. For practical purposes it might in most cases be considered continuous.)

Another choice is between giving an *absolute* or *relative* measure of punctuality. Here, the amount of (un)punctuality is not, respectively is, set in relation to production volume, e.g. scheduled travel time. E.g. the percentage of passengers arriving to their destination within 15 minutes after scheduled time is a relative measure, while the sum of their minutes after schedule is an absolute measure.

Another choice *is where to measure*: underway (continuously or at certain locations, e.g. stations), at origin and/or destination. We now describe how a choice here, might have different implications dependant on choice of system boundaries. Consider the choice of destination punctuality. To a passenger travelling from A to B this is clearly the punctuality of arrival at B. To the train pursuing the distance A-B-C, arrival punctuality at B is one of the train's underway punctualities and the time of arrival at C is the train's destination punctuality. If only the vehicle's destination punctuality is recorded, it is impossible to tell passenger's destination punctuality.

Assessment (värdering) principles of punctuality. There are several alternative assessment principles of (un)punctuality. The principle chosen is not always explicitly stated. E.g. by choosing to measure the total number of delay minutes to passengers, the tacit assessment is made that each delay minute of each passenger is equally important. The choice of assessment principle partly follows from choices mentioned earlier, as might be seen from the following example. The punctuality is considered as the proportion of passengers arriving within 15 minutes from schedule. This implies customer system boundary, discrete punctuality and relative punctuality. Due to discrete punctuality, the value of time per passenger is non-linear (as times over 15 minutes per passenger is not accounted for) but the linear to the number of passengers. So the choice of assessment function falls into either of the categories

- Linear to unpunctuality time
- Nonlinear to unpunctuality time

A choice is also to be made between

- Linear to people (goods) affected
- Nonlinear to people (goods) affected

A choice of assessment principle is whether to give different values to

- Unanticipated unpunctuality and
- anticipated unpunctuality

A rationale to assess the former higher is that such disturbances are more difficult for passengers to cope with. A rationale against is that notoriously unpunctual operators come out better from a given disturbance than usually punctual operators (at least in a relative sense). The extreme of anticipated unpunctuality is of course to

- prolong the scheduled travel time

in order deal with uncertainties in travel time.

The total cost of travel time and unpunctual time could be minimised when we know their respective costs, e.g. per hour late. However, such a minimisation might also take into account the risk of running late, as it not clear-cut that e.g. a certain delay of 10 min is assessed as equal to a 50% probability of running 20 minutes late. So, the *risk-taking profile* is important.

Another choice is whose costs to assess, e.g.:

- Customer's
- Vehicle operator's
- Infrastructure operator's
- Social economy costs.

The assessment of *cancelled* or *partly cancelled trains* is a tricky question and is not treated here.

Definition. By *regularity* (regularitet) is meant to what extent a train adheres to the scheduled travel time along given distance/s.■

The antonym to regularity is irregularity (oregularitet). The definition of *regularity* must be complemented by the distance/s concerned. A train with an irregularity of 0 on the distance AC, might have a irregularity of +2 minutes on AB and -2 minutes on BC. Hence it might be worth introducing *absolute irregularity* to give a measure of deviation from schedule. In this case it amounts to 4 minutes, but note that a change in the location of B might change this number.

Many of the definitions so far have assumed the existence of a given train schedule. However, the capability of the train traffic process, i.e. how little travel times vary, is an important input to train traffic planners. It might also be important to e.g. passengers. We give just a simple example of a measure of this *variation*:

Standard deviation of transport time / Distance of transport.